THE RISE OF Artificial Intelligence

Real-world Applications for Revenue and Margin Growth

Zbigniew Michalewicz
Leonardo Arantes
Matt Michalewicz

Published by Hybrid Publishers
Melbourne Victoria Australia

Hybrid Publishers,
PO Box 52, Ormond VIC Australia 3204.
www.hybridpublishers.com.au

NATIONAL LIBRARY OF AUSTRALIA
A catalogue record for this book is available from the National Library of Australia

First published 2021
Text design & layout by Midland Typesetters, Australia
Cover design by Marchese Design
ISBN 9781925736625 (p)
Printed by Tingleman Pty Ltd

Dedicated to the entrepreneurs, scientists, and business leaders that have paved the way for Artificial Intelligence over the decades past, and are paving the way for its future in the decades to come.

PREFACE

What This Book is About and How to Read It

"We're at the beginning of a golden age of AI. Recent advancements have already led to inventions that previously lived in the realm of science fiction—and we've only scratched the surface of what's possible."

Jeff Bezos, *Amazon CEO*

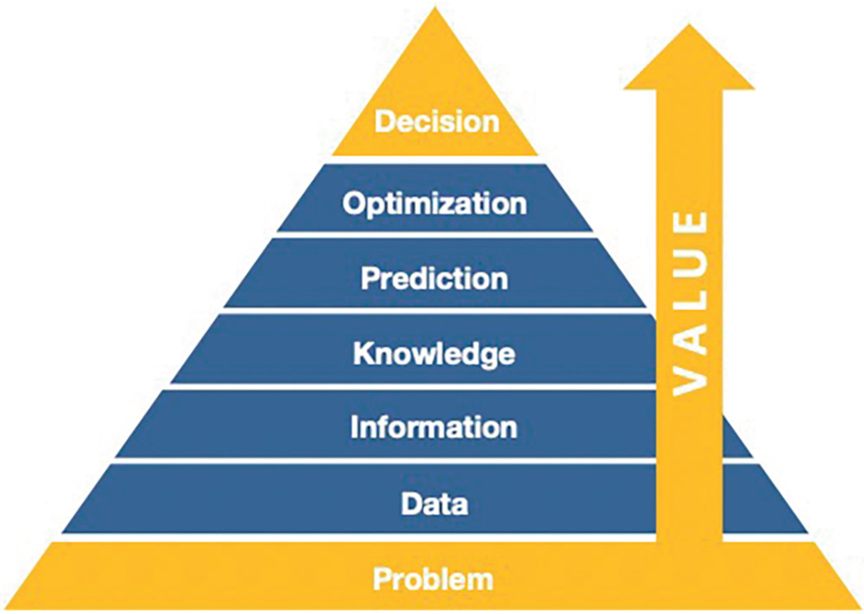
Few terms have captured our imagination in recent times like "Artificial Intelligence." And not just through sensationalized media articles about how AI will soon displace all jobs and rule the world, but also through movies, books, and television shows. It now seems that everyone "knows" about AI; that everyone has an opinion. And yet, in our experience, few people actually understand what Artificial Intelligence is and isn't, where the field came from and where it's heading, and how the technology can be harnessed to generate commercial outcomes.

Given the immense amount of disinformation and misunderstanding, we have written this book to demystify the subject of AI and explain it in simple language. Most importantly, we have written this book with the business manager in mind, someone interested in the topic from a real-world, commercial perspective—a perspective of how the technology can create value and increase competitiveness *today*, rather than what might happen in 25 years' time or how a superior intelligence might overcome the human race in the distant future. Such philosophical treatises are thought-provoking (to say the least) and the subject of many books published each year, but this isn't one of them. Instead, *The Rise of Artificial Intelligence* provides a commercial exploration of AI, with particular emphasis on how AI-based systems can improve decision making in organizations of all shapes and sizes.

As such, this book presents Artificial Intelligence through the lens of decision making for two reasons: First, because the world has reached a level of such unprecedented speed, complexity, and noise, that no one can assess and evaluate all the available data when making decisions; and secondly, because the decisions we make affect the outcomes we achieve. In other words, better business decisions lead to better business outcomes. Although Artificial

Intelligence can be applied to many areas besides decision making—such as automation and robotics, or image and speech recognition—these subjects don't feature heavily in the pages ahead except for Chapter 1, where we provide an overview of the research areas of AI. Ultimately, revenue and margin growth comes down to the decisions an organization makes (or doesn't make), and hence the application of AI to decision making is our primary focus.

To best present the concepts in this book, we've used a *problem-to-decision pyramid* to represent the continuum that exists in terms of an organization's ability to improve its decision making:



Each layer of this pyramid represents a step in the journey for improved decision making: the higher we go, the better our decisions (and the more value we can create). The structure of *The Rise of Artificial Intelligence* reflects the structure of this pyramid, with the first two parts of the book investigating each layer of the pyramid, and the last two parts illustrating the application of Artificial Intelligence to real-world problems for the purpose of generating revenue and margin growth.

Chapter 1 begins with a high-level overview of Artificial Intelligence—its history, areas of research, and current progress and challenges—before introducing the *problem-to-decision pyramid* in Chapter 2, which conceptualizes the journey from defining a problem to making a decision through the use of data, information, knowledge, prediction, and optimization. Chapter 3 concludes Part 1with an in-depth examination of a complex business problem set in the fast-moving consumer goods industry, which is used to explain the role of objectives, business rules and constraints, and the application of Artificial Intelligence algorithms for improved decision making.

This complex business problem of promotional planning and pricing is then used as a running example throughout Part II, which explores the inner workings of predictive models, optimization methods, and various learning algorithms. Because data and modeling form the basis of prediction and optimization, this part of the book opens with a chapter on data and modeling, along with a discussion of common issues such as data availability, completeness, and preparation. In Chapters 5 and 6 we review various AI and non-AI methods for predictive modeling and optimization, whereas in Chapter 7 we present adaptability and learning concepts—which together (i.e. prediction, optimization, and self-learning) comprise the backbone of any AI-based software system.

As an important aside, Chapters 4 through 7 represent the most technical material of the entire book, attempting to explain the innermost mechanics of several Artificial Intelligence algorithms such as neural networks and genetic programming. Although non-technical readers can easily progress through Part II to gain a deeper understanding of algorithms and models, readers without an interest in data, problem modeling, or how Artificial Intelligence algorithms work, can jump straight to Part III, which presents real-world applications of Artificial Intelligence.

The application areas in Part III explore the problem-to-decision pyramid in the context of real-world problems and business objectives, covering both the lower layers of the pyramid focusing on data and the analytical landscape of an organization (i.e. information and knowledge), as well as the upper layers of prediction, optimization, and self-learning, and how they're enabled by Artificial Intelligence methods. For ease of reading, we've divided Part III into three chapters, each being dedicated to a specific business function—in particular, *sales, marketing,* and *supply chain*. These case studies are based on an enterprise software platform called Decision Cloud®, which is a modularized, cloud-based platform that empowers staff to make better and faster decisions through the use of Artificial Intelligence.

And finally, Part IV concludes the book with common questions and concerns that organizations have on the application of Artificial Intelligence, such as: *"Would AI work for me?"* and *"Where should I start?"* These two chapters provide practical advice for selecting the right business problem, developing a business case, choosing a technology partner, as well as other topics such as digitalization and change management.

To improve the reader's understanding of the content, we've also created a set of supplementary videos that can be accessed at: www.Complexica.com/book/RiseofAI/. These videos bring to life the concepts presented in each chapter—for example, by providing a visual explanation of ant system algorithms in Chapter 1, the layers of the problem-to-decision pyramid in

Chapter 2, the workflow of promotional planning and pricing in Chapter 3, and so on. In these videos we're able to "show" concepts that can only be "told" within the confines of the printed page.

In terms of how to read this book or watch the videos, the ideal way is to progress sequentially from Chapter 1 to 12. For the less technically-inclined reader, however it's possible to jump around in any sequence that best satisfies curiosity and interest. For example, the reader might begin with an overview of Artificial Intelligence in Chapter 1, then progress to the application areas in Chapters 8, 9, and 10, before returning to Chapters 2 and 3 to better appreciate the problem-to-decision pyramid and the intricacies of solving complex business problems (after all, why are complex business problem so difficult to solve?). Alternatively, a reader might start with the application areas in Chapters 8, 9, and 10, then move back into Part II to better understand how algorithms and models work, before progressing to Part IV for practical advice for initiating an Artificial Intelligence project.

However, regardless of the reader's technical sophistication or their interest in the implementation aspects of AI-based software, it's highly recommended that everyone start with the first two chapters for an introduction into the world of Artificial Intelligence and an overview of basic concepts and terminology. From this perspective, the sequence of reading the remaining chapters is of far lesser importance.

Lastly, we'd like to say that the material presented in this book is the result of 40 years of first-hand Artificial Intelligence research within university settings, and more than twenty years of implementing AI-based enterprise software systems in many (often very large[1]) organizations across three continents. With that in mind, we'd like to thank everyone who made this book possible, with our special appreciation going to many Australian companies we collaborated with over the years in the application of Artificial Intelligence, such as PFD Foods, BHP Billiton, BMA, Pernod Ricard Winemakers, Lion Drinks, Bunzl, DuluxGroup, Rio Tinto, Metcash, Pfizer, Janssen, Haircare Australia, Fortescue Metals Group, CBH Group, Roy Hill, Glencore, Polyaire, Treasury Wine Estates, and Costa Group. Within these companies, we'd like to thank Chris Baddock, John Barakat, Renato Bellon, Simon Bennett, Damian Bourne, Warren Brodie, Michael Brooks, Pierre-Yves Calloc'h, Daryl Chim,

1 Our experiences of implementing enterprise-grade software based on the latest Artificial Intelligence algorithms and methods are based on many projects with global giants—such as BHP Billiton, General Motors, Bank of America, Pernod Ricard, Unilever, Air Liquide, Ford Motor Company, Glencore, Beiersdorf, Rio Tinto, and ChevronTexaco, among many others—as well as smaller companies that benefited from the research & development and innovation carried out by these larger organizations.

Richard Cohen, Jevan Dickinson, Andrew Endicott, Eglantine Etiemble, Scott Fellingham, Greg Feutrill, Garth Gauvin, Ward Gauvin, Scott Graham, Chris Green, Kylie Grigg, Richard Hansen, Mark Hayden, Kim Heatherton, Mark Ivory, James Jones, Mike Lomman, Brett McKinnon, Stuart McNab, Doug Misener, Luke Mitchell, Stephen Mooney, Aemel Nordin, Mark Powell, Rod Pritchard, Robin Pyne, Mathew Regan, Darryl Schafferius, Mark Shephard, Jon Simpson, Kerry Smith, Richard Taylor, Soner Teknikeller, Lance Ward, John Warda, and Joel Zamek.

We'd also like thank a few individuals who contributed to the content and ideas in this book, namely, Reza Bonyadi, Łukasz Brocki, Tom Heyworth, Xiang Li, Łukasz Olech, Ali Shemshadi, Larisa Stamova, Chris Zhu, as well as members of Complexica's scientific advisory board who we've worked with over the years: Reza Bonyadi, Łukasz Brocki, Longbing Cao, Raymond Chiong, Vic Ciesielski, Carlos Coello, Ernesto Costa, Kalyanmoy Deb, Kenneth De Jong, A.E. Eiben, Xiaodong Li, Masoud Mohammadian, Pablo Moscato, Frank Neumann, Zbigniew Raś, Markus Wagner, Thomas Weise, Adam Wierzbicki, and Mengjie Zhang.

And finally, it was a great pleasure to write about a topic that's been the central focus of our working lives for so many years, and we hope that readers enjoy this book as much as we enjoyed writing it. We believe that anyone in any organization who makes operational, tactical, or strategic decisions—whether on the factory floor or in the boardroom—will find this book valuable for understanding the science and technology behind better decisions. Enjoy!

Adelaide, Australia
March 2021

Zbigniew Michalewicz
Leonardo Arantes
Matt Michalewicz

TABLE OF CONTENTS

PART I

Artificial Intelligence as Applied to Decision Making

CHAPTER 3

An Extended Example: Promotional Planning and Pricing

"Half the money I spend on advertising is wasted;
the trouble is I don't know which half."

John Wanamaker, *Merchant and Politician*

Having defined the problem-to-decision pyramid—along with each layer and how it contributes to the decision-making process—let's now turn our attention to the challenge of *promotional planning and pricing*[1] in the fast-moving consumer goods (FMCG) industry, which has all the hallmarks of a complex business problem worthy of discussion.

We've all experienced product promotions *(Sale! 50% off! Buy one, get one free!)*, which manufacturers and retailers use to drive foot traffic into stores, increase volume and market share, and build awareness for new products (have you ever bought a product on promotion, liked it, and then switched to that product permanently?). These promotional activities are typically funded by both the retailer and participating manufacturer, and can account for almost 20% of the revenue of FMCG companies. Hence, promotions are a big deal. And not only is the problem of promotional planning inherently complex—where it's difficult to make "good" decisions that generate real improvements in revenue, margin, and overall profitability—but it's also an inherently *high-value problem,* where the difference between good and bad decisions can mean millions of dollars and many percentage points of market share.

Now, suppose we work for a large, global manufacturer, one that produces many household products—ranging from snacks such as ice-cream and peanut butter, through to common toiletries like toothpaste—and most of our sales come through retailers, especially supermarkets, where promotions drive almost 80% of our sales volume. Within this setting, let's say we're in charge of promotional planning for all product categories sold within a single retail

1 There are several related terms to the area of *promotional planning*, including *promotional programming*, *trade promotions*, *trade promotion optimization (TPO)*, and so on. For the sake of simplicity, however, we'll use "promotional planning" throughout this text even though at times there may be a more fitting term.

chain, called Mary's Market, a fictitious retailer in Australia with 500 stores across the country. We'd like to create the "best" possible plan for Mary's Market (i.e. one that achieves the best result for the objectives we're trying to minimize or maximize), but this isn't easy given the complex nature of the problem and the fact that our objectives (e.g. maximizing profit) can often be in conflict with those of Mary's Market—for example, if they request a deeper discount for an upcoming promotion, it might erode our profit if the volume uplift isn't great enough.

In fact, we can go further on this point by saying that Mary's Market (and retailers in general) are working primarily to build their own sustained profitable growth and are ultimately indifferent to an individual supplier's growth or profit concerns. This creates a critical bottleneck, in that as the manufacturer, we can only "grow through" the retailer (at least for this channel). Hence, in addition to the science of promotional planning and pricing (of finding the "best" possible plan), there's also the "art" of negotiating the final plan with Mary's Market so it can be executed in their stores, rather than being rejected or significantly modified. In this chapter we'll focus on the scientific aspects of the promotional planning problem, while acknowledging that being able to negotiate well with retailers requires a sound understanding of the consumer, shopper, retailer, and the overall category (as discussed further in the Trade Spend Optimization case study in Chapter 9.4).

Basic Terminology

Before exploring this problem in detail, let's first cover some basic terminology in the context of promotional planning:

- *Store*: is an individual retail store that might either be part of a chain or independently owned.
- *Retail chain*: is a network of stores owned by the same company.
- *Banner group*: is a network of independently owned stores that sign up to a collective "banner" (i.e. brand name) to leverage economies of scale in procurement and marketing.
- *Geographic boundaries*: is the geographical region for which a promotional plan is made for a particular retailer, which is typically a state (e.g. VIC or NSW in Australia) but could also be more granular, such as metropolitan areas or rural regions.
- *Planning horizon*: refers to how far out we plan our promotions, which could be yearly, half-yearly, quarterly, or some other time horizon.
- *Promotional period*: is the length of time the promotion runs. For example, a promotional period may be one week for a grocery retail

chain, which means that a plan needs to be created for each week of the year. In other industries, the promotional period might be fortnightly or even longer.

- *Promotional plan*: is the set of all promotional activities during the planning horizon (also called the *promotional calendar*).
- *Variable*: is any value that can be changed when we search for the best plan. For example, variables may include the promotional price, depth of discount, and the promotion type, among many others.
- *Predictive model*: provides a sales forecast for each product within each promotional period. A predictive model for promotional planning may also include external data, such as weather, sporting events, or public holidays.
- *Promotion type*:[2] refers to the type of promotion, such as "% off," set dollar discounts, buy-one-get-one-free (BOGO), or multi-buy (i.e. multiple products offered for a fixed price).
- *Price step*: refers to the minimum amount that a price can be discounted. For example, the discounting of a $19.95 shelf price might occur through 50 cent *price step*s.
- *Objectives*: are measurable business goals we want to minimize or maximize. For promotional planning, typical objectives include maximizing volume and gross profit.
- *Business rules*: are something we might choose to do because there are valid business reasons to do so—for example, we might have business rules for minimum and maximum promotional frequencies for specific products, along with a minimum sell price or margin.
- *Constraints*: on the other hand, are something we're obligated to do, usually because of commercial commitments (such as trading terms), government regulations, capacity limitations, or some other factor. Both business rules and constraints might be "hard" (meaning they cannot be violated), or "soft" (meaning they can be violated in order to achieve a better overall result), and may apply to individual products and/or the overall plan (such as a minimum KPI or business objective that needs to be achieved for the plan to be acceptable).[3]

2 Also commonly referred to as the *promotional mechanics*.

3 Even though constraints represent something we're obligated to do (usually because of commercial agreements), it's worth exploring whether the violation of some constraints might produce a significantly better result (and if so, potentially lead us to renegotiate a commercial obligation in order to execute the better plan). For this reason, some constraints are defined as "soft" in order to test if they're detrimental to achieving superior business outcomes.

- *Feasibility and infeasibility*: any promotional plan we create is either *feasible* if it doesn't violate any hard business rule or constraint, or *infeasible* if it does violate a hard business rule or constraint.

3.1 The Problem: Promotional Planning in FMCG

Recall that we work for a global manufacturer that sells household products—including snacks and toiletries—through major retail chains, and we're in charge of promotional planning for one of these retail chains, Mary's Market in Australia. However, it's important to point out that our company has a very limited route to market, meaning that we can only sell our products through a small number of retail chains, with Mary's Market being one of them. Whereas Mary's Market, on the other hand, has an abundant choice of products not only from us, but from all our competitors. As an example, we might offer 32 different ice-cream products and eleven peanut butter products, and Mary's Market is only one of fifteen retail chains through which we can sell these products. Mary's Market, however, has access to hundreds of different ice-cream and peanut butter products from all the various manufacturers that produce those products. This raises the stakes for us on the importance of promotional planning, because there's a limited amount of shelf space available for which many manufacturers are vying, and so if our promotions aren't successful, then Mary's Market might delete our products from the shelf in favor of those offered by our competitors.

Now, let's consider the structural elements of the problem: In addition to the products we offer, Mary's Market also sells many other product categories we don't supply, like fresh fruits and vegetables, clothing, magazines, and pet care products, among others. For the product categories relevant to us, there's a certain length of time for which we must plan our promotions, and let's assume that we plan for the entire year, so our planning horizon is 52 weeks. Let's also assume that the promotional period is one week, which provides us with the plan's granularity and the level of detail we must plan to. We can represent this granularity with a promotional calendar—or *slotting board*—where each column is a week, and each row is a particular product, allowing for promotions to be *slotted* into each column/row combination. For example, the below slotting board is for the snacks product category in NSW:

Retailer:	Mary's Market
State:	NSW
Category:	Snacks

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	...	WK 51	WK 52
Product 1									
Product 2									
Product 3									
Product 4									
Product 5									
Product 6									
...									
Product 100									

If there are 100 individual products for us to plan in the snacks category, then the simple decision of whether or not to promote a particular product for any given week requires 5,200 individual "yes" or "no" decisions (52 weeks × 100 products). If we ignore other elements of the problem—such the promotional price, promotion type, ancillary marketing, holidays and seasonality, catalogue placement, and so on—the amount of individual binary yes/no decisions still implies an astronomical number of possible plans (1 followed by 1,565 zeros!), each of which would look something like this on the slotting board:

Retailer:	Mary's Market
State:	NSW
Category:	Snacks

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	...	WK 51	WK 52
Product 1	Y	Y		Y	Y				
Product 2	Y		Y		Y	Y			Y
Product 3	Y		Y		Y	Y			Y
Product 4			Y						
Product 5	Y			Y					
Product 6	Y	Y		Y					
...									
Product 100	Y			Y	Y			Y	Y

Our promotional plan must also adhere to business rules for individual products. For example, business rules may prevent the promotion of specific products for less than four weeks or more than twelve weeks during any twelve-month period, known as *minimum and maximum frequency*. Furthermore, these business rules may apply to the "gap" in between promotions for the same product, known as the *minimum and maximum promotional gap*, so that promotions don't happen too often or too infrequently (such as promoting the same product for nine consecutive weeks and then not promoting it for the rest of the year). Such business rules also apply to prices, where the minimum promotional price might be set at no less than 50% of the shelf price (i.e. the non-promoted price) and not more than 90%, and must move by some price step increment (e.g. 50 cents or one dollar, to avoid awkward prices like $8.13). These business rules can be visualized in the following table:

	Shelf Price	Min Freq	Max Freq	Max Promo Length	Min Promo Gap	Max Promo Gap	Min Promo Price	Max Promo Price	Price Step	Allow rounding?
Product 1	$23.50	13	26	2	2	6	$12.00	$20.00	$0.50	Y
Product 2	$14.50	13	26	2	2	6	$10.00	$12.50	$0.50	Y
Product 3	$19.00	13	26	2	2	6	$10.00	$15.00	$1.00	N
Product 4	$22.00	20	39	3	2	4	$11.00	$18.00	$1.00	Y
Product 5	$52.00	20	39	3	2	4	$40.00	$48.00	$1.00	Y
Product 6	$48.00	13	39	3	2	4	$40.00	$45.00	$0.50	Y
...										
Product 100	$65.00	13	13	1	2	6	$55.00	$60.00	$1.00	Y

We need to also bear in mind that different geographic regions may have their own business rules, which adds further complexity to the problem. There may also be specific rules tied to the promotional period itself, such as the minimum and maximum number of products on promotion at any given time (which may differ from week to week as we consider holidays or other events). For example, during most weeks of the year, it may be permissible to have 30% of our products on promotion at the same time, but for certain weeks of the year, such as before major holidays like Easter, Christmas, and New Year's, it may be appropriate to increase this percentage.

These business rules may be quite complex, and yet, we've only considered products in isolation, and not thought about products being constrained by the promotional activity of other products. Hence, we might need to extend our business rules to cover different pack sizes of the same base product (e.g. 45 gram- and 170 gram-bags of the same snack), different varieties or flavors of the same product (e.g. where all go on promotion or none at all), or different products altogether (e.g. where we can promote a subset of our snacks at any given time, but not all of them within the same promotional period). And lastly, the promotional plan as a whole must meet certain KPI thresholds, such as volume or volume growth, revenue or revenue growth, and gross profit (among others), which we must also define as business rules.

Keeping all this in mind, let's turn our attention to the slotting board for Mary's Market in NSW (for the snacks category), which may look like the following:

Retailer:	Mary's Market
State:	NSW
Category:	Snacks

Promotion Type	In store
Promotional Price	$19.00
Shelf Price	$23.50
Discount	$4.50
Margin	35%
Customer Margin	33%
Min/Max Frequencies	13/26 Weeks
Min/Max Sell Price	$12/20
Min/Max Promo Gap	2/6 Weeks

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	...	WK 51	WK 52
Product 1	Y	y		y	Y				
Product 2	y		y		y				y
Product 3	Y		y		y	Y			y
Product 4			Y			y			
Product 5	Y			Y					
Product 6	Y	y		y					
...									
Product 100	Y			Y	Y			Y	Y

We can see that beyond the simple yes/no decision of whether a product is on promotion during any given promotional period, each cell has a set of values associated with it. These are color coded in the table above on the right, and broken down as follows:

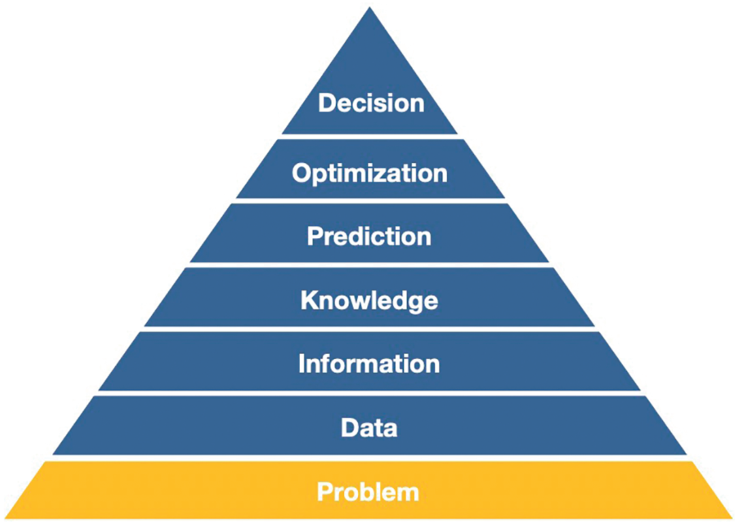
- *Variables* (*green*): are values we can change and "trial" when searching for the best plan. In the right-hand side table above, these include the promotional price and promotion type, which are often dependant (or linked) because particular promotion types may be associated with particular price points. For example, a "%-off" promotion will always link the promotional price with a corresponding percentage discount level.
- *Reference points* (*yellow*): are values we don't adjust when creating new promotional plans. In the right-hand side table above, this is the shelf price (i.e. the non-promotional price), which might change only once or twice a year whenever there's an overall price change.
- *Derived values* (*orange*): are calculated from other values and may sometimes act as business rules and constraints. In the right-hand side table above, these include the discount, margin, and retailer margin, which are all calculated from the shelf price, promotional price, and overall product cost.
- *Business rules* and *constraints* (*blue*): represent limits imposed on particular values for the purpose of optimization. In the right-hand side table above, these include the minimum and maximum promotional price, the minimum and maximum promotional frequency over any given planning horizon, and the minimum and maximum gap between promotions for the same product.

When we consider all these additional factors, it seems that our job is quite challenging! Just creating a single feasible plan—one that doesn't violate any business rules or constraints—is already a Herculean task, most likely involving a multitude of spreadsheets and endless hours of evaluating new plans against historical promotions to "guess" the likely outcome. Then, on top of all that, the plan won't be executed and produce the desired upside unless we secure buy in from Mary's Market. This means we need to create a "win-win-win" slotting board—one that optimizes marketplace outcomes for ourselves, Mary's Market, as well as for the overall category—a slotting board that represents a needle in an astronomical haystack!

For more information about the myriad complexities of promotional planning and pricing, as well as the use of rules and constraints to create slotting boards, we encourage you to watch the supplementary video for this chapter at: www.Complexica.com/RiseofAI/Chapter3.

3.2 Applying the Problem-to-Decision Pyramid

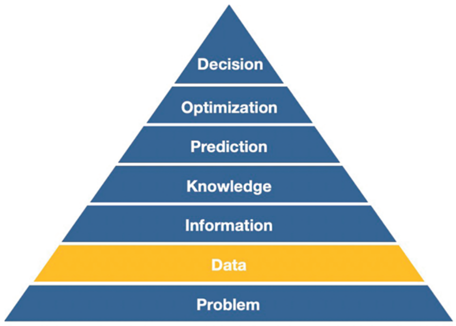
Considering the difficulty of creating a single feasible plan, what would it take for us to create an optimal plan? And by what measure, or measures, should the plan be evaluated to determine its "optimality"? Before we answer these questions, let's revisit the problem-to-decision pyramid from Chapter 2, and review these layers in the context of promotional planning:



Given that we've already described the problem of promotional planning in some detail, let's move to the next layer of the pyramid: *data*.

Data

Data is the fundamental "raw material" for promotional planning, as any decision we make will be dependent on the underlying accuracy of our predictions, which, in turn, will be dependent on the available data. In the context of promotional planning, what we're most interested in is data that can help



us infer shopper behavior, particularly consumer demand—after all, promotional activities are usually undertaken to increase demand. Hence, we're particularly interested in *point of sale* data, which is the data closest to the consumer (often called *sell out* data), as opposed to the retailer's orders from a manufacturer (often called *sell in* data), which may be subject to various supply chain related factors such as inventory policies and minimum order quantities. When such data is overlaid with promotional activities and pricing history, we can infer long-term trends at the product or category level, as well as seasonality by examining the baseline (non-promotional) sales volume over time.

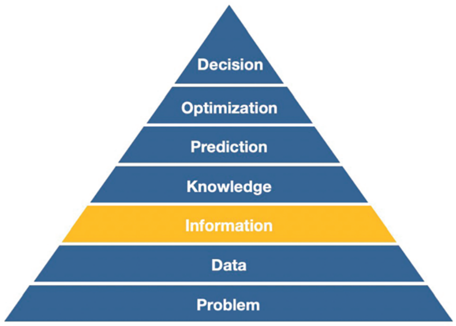
It's common knowledge that "garbage in equals garbage out," thus, data quality and availability are important considerations. Unfortunately, most modern organizations struggle with a variety of "data issues" such as:

- *Missing data*: where a critical piece of data has been overwritten or wasn't collected in the first place. A classic example is stock on hand data, where most inventory management systems only contain the most current stock on hand without retaining any historical values. However, this historical data is useful for identifying stockouts in the past. When overlaid with sales data, it may become evident that some unexplained drops in demand were actually due to stockouts rather than actual drops in demand. Another example is when a retailer doesn't provide detailed point of sale data (by store, by day, by product), which compromises a manufacturer's ability to build an accurate prediction model for promotional planning.
- *Dirty data*: which typically stems from master data management issues within IT systems (such as ERP or point of sale). Examples of dirty data include products assigned to the wrong product category, or the use of free text fields to describe product information such as pack size and volume. Furthermore, using multiple point of sale systems may result in sales data that's in multiple formats, requiring considerable time and effort to standardize.
- *Incorrect mappings*: occur when some data mappings change over time. A typical example within promotional planning is the mapping of individual stores to a particular retail chain. In some industries, stores are independently owned and may switch from one retail chain banner to another. If these moves aren't recorded over time, then an

incorrect data mapping occurs, which may lead to historical demand not being correctly recorded for the same store. Another example is when new products are launched that replace existing products. For example, a 500g product is discontinued, and 350g and 650g products are launched within the same product line. New product introductions and even simple pack size changes might be incorrectly mapped within IT systems, leading to issues when that data is used to create a predictive model.

Such data issues are typical within most modern organizations and must be dealt with as we climb the problem-to-decision pyramid. Fortunately, there are a variety of methods we can use, which we'll discuss in the following chapter on data and modeling.

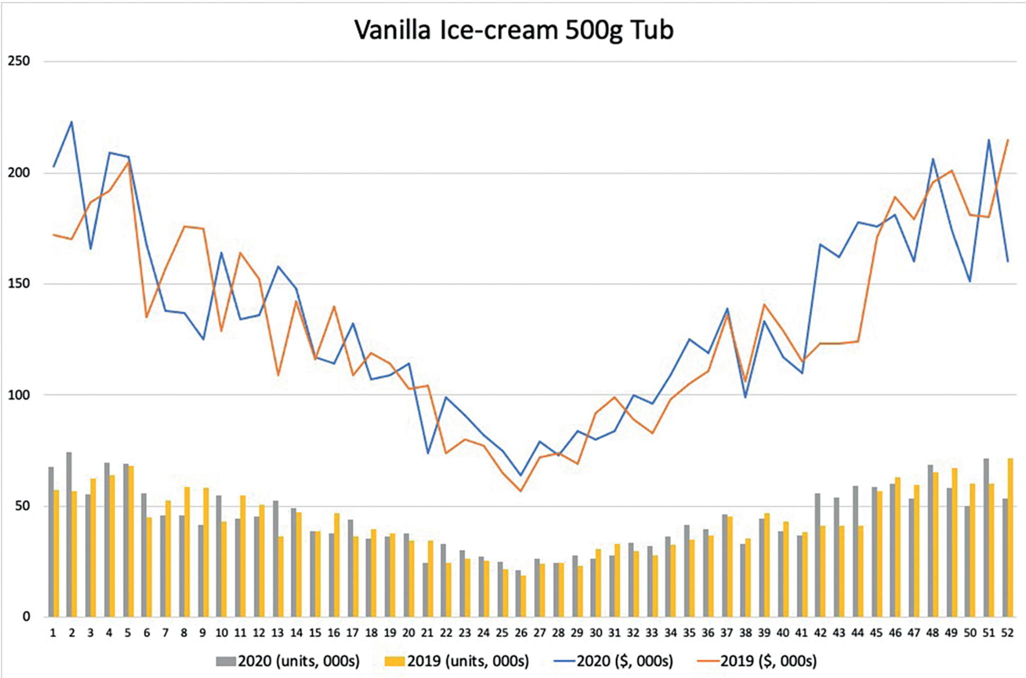
Information



Most organizations are overly dependent on informational reports, especially within business functions characterized by complex planning and scheduling activities. Promotional planning is no exception, with spreadsheets and pivot tables often being the norm for pulling data into tables and charts. Some typical examples of reporting within the context of promotional planning include:

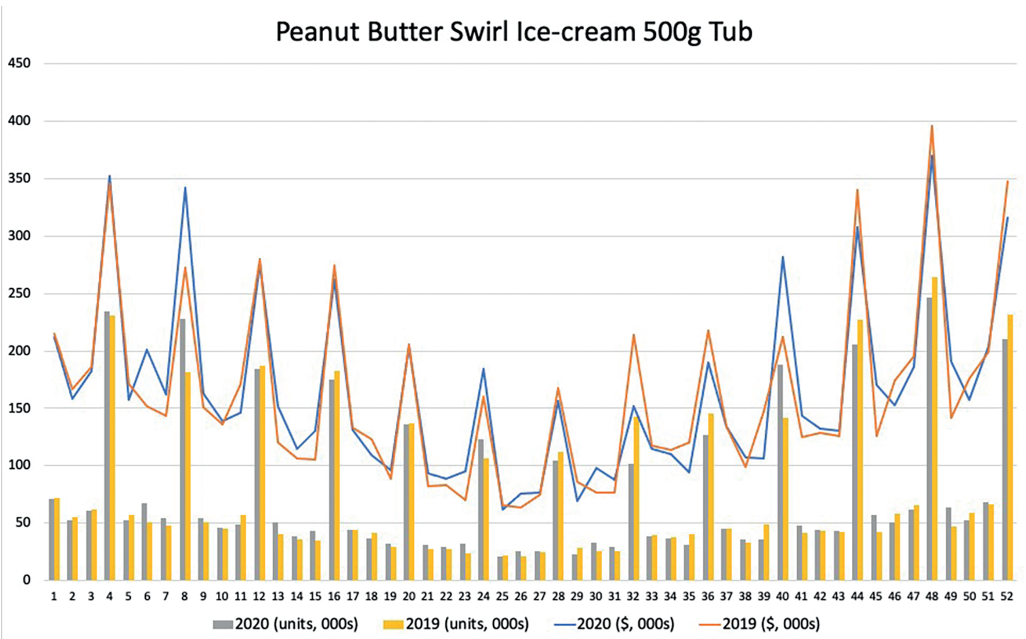
- Sales data in $ (gross revenue)
- Sales data in $ (net revenue)
- Promotions and pricing data
- Volumes sold (units and/or kgs)
- Gross profit
- Trade spend
- Trade spend as a percentage of gross revenue

The above information may be viewed in various ways: *per promotional period, per store, per banner, per state, per product category, per brand, per product,* or any combination thereof. A typical example of data visualization is shown below, where two years of sales data are displayed in units and revenue:



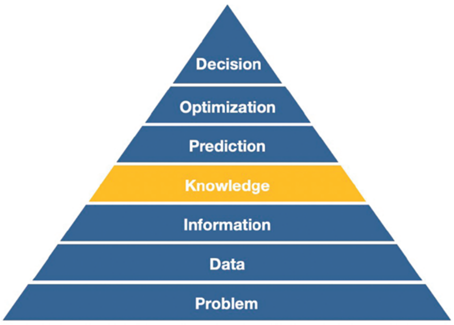
The chart shows a product where no price changes or promotions have occurred over the two-year period, so revenue for any given week is three times the units sold. The steady U-shape movement of sales going down and then up is the *seasonality effect*, with higher sales occurring in summer rather than winter (southern hemisphere data). But beyond this immediately evident seasonality effect, the remaining variability (i.e. peaks and troughs) are not easily explained, requiring further data analysis and investigation.

If we consider another snack product, but one that's promoted at 50% off every four weeks, the chart will look very different:



Note the pronounced volume and revenue peaks from the promotional activity every four weeks, along with the seasonality effect from summer into winter and then back again. Such reporting serves as the bare minimum for assessing the effectiveness of promotions and conducting fundamental financial analysis, but is inadequate for optimizing our decisions about upcoming promotions.

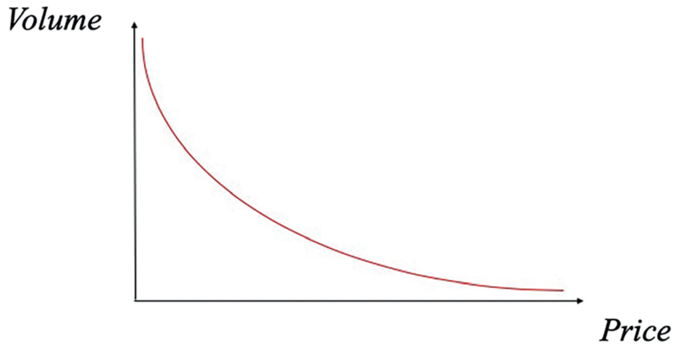
Knowledge



Knowledge is typically generated through some deeper analysis of historical data, and attempts to answer the question: *Why did this happen*? Such analysis usually involves the overlaying of several datasets, which can then be visualized in a chart or graph. For example, viewing sales data for a particular product, overlayed with promotional periods and known price changes, may reveal peaks and troughs associated with promotional activity, seasonality, and other factors.

Let's discuss this layer of the pyramid in the context of an everyday product, such as toothpaste. This product is generally used twice a day, every day, in almost the same amount. Not much affects the quantity used by consumers, as the overall demand for toothpaste is relatively fixed, and doesn't fluctuate with time of year, weather, public holidays, or other external factors. However, there are many varieties of toothpaste: regular toothpaste, anti-plaque, anti-calculus, antimicrobial, sensitive, whitening, and children's toothpaste. Most of these sub-categories have basic, mid-range, and premium versions, with some sub-categories even having super-premium versions (such as sensitive and whitening). Furthermore, multiple brands might cover most or all of these niches, and the same product might be available in different sizes.

Using data related to historical toothpaste sales, we can investigate many factors that affect demand, the most important of which is *price elasticity* (or simply *elasticity*). Elasticity has its roots in Economic Theory and is part of the Law of Demand, which states that demand for any given product will go up as its price goes down, and vice versa (with a few exceptions, such as some luxury goods and other limited circumstances). All products have an elasticity curve, such as the one shown below. Some of these curves are steep, where a small change in price causes a large change in demand, and some shallow, where a large change in price causes a small change in demand:



Data related to historical toothpaste prices can help us plot an elasticity curve that shows how permanent changes in price affect demand, while the historical performance of promotions at different price points can help us create a *discount elasticity* curve that shows how temporary (i.e. promotional) changes in price affect demand. Since consumers are much more sensitive to a temporary discount in price than a permanent change, discount elasticity is usually much more pronounced than price elasticity for the same product.

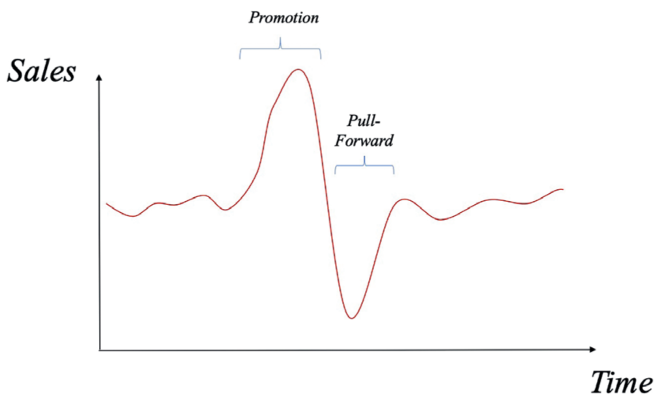
For example, if the toothpaste has a shelf price of $6.50 and has participated in regular “20% off” promotions, as well as less frequent “50% off” promotions, then historical sales data would almost certainly reveal significant volume increases during the “20% off” promotions and even greater volume increases during the “50% off” promotions. With enough historical data of actual promotions, we can plot a discount elasticity curve and extrapolate what the sales volume increase might be for a “30% off” promotion, even though we’ve never run that exact promotion before.[4]

We can also increase our knowledge by trying to understand *why* a reduction in price caused an increase in sales. For some products, a change in price may have affected a consumer’s decision on whether or not to buy the product at all, so the promotion resulted in a real change in consumption. This isn’t the case with toothpaste, however, as people tend to brush their teeth twice a day regardless of whether toothpaste is on promotion or not—the only difference being the brand and type used. Therefore, the real consumption of toothpaste doesn’t change during promotional periods, but rather, consumers switch between products and this leads to higher sales of one toothpaste over another. However, some products such as ice-cream and champagne may experience a real increase in consumption when placed on promotion, because people not only buy more ice-cream and champagne, but they also consume more as well.

4 Such extrapolations are useful until we “overdo it” by promoting too frequently, which may lead to a permanent change of perception for that product’s value and when consumers buy it (e.g. they never pay more than $5 for that product, so they’ll wait and stock up during the next promotion).

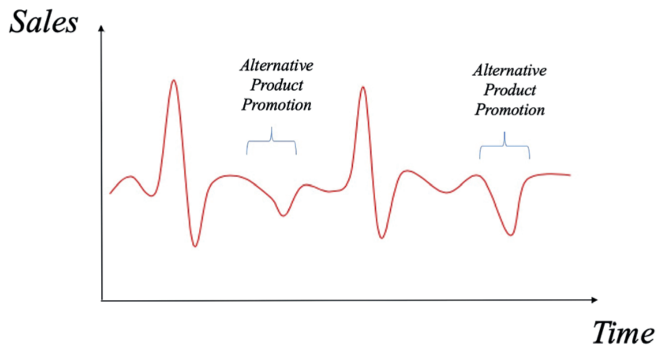
Some products are highly seasonal, like hay fever medicines, and we may find seasonality to be the dominant factor driving sales (with our analysis possibly revealing that any increase in sales wasn't attributable to the promotion). Furthermore, we may discover long-term trends for particular products, brands, or even entire categories (as some categories might be experiencing growth while others are in a state of decline), providing us with even more knowledge of why demand increased or decreased at certain points in time.

And lastly, when toothpaste is on promotion, we can expect consumers to stock up on the product, therefore "bringing forward" future purchases. This is called the *pull-forward effect*, which influences non-perishable products (those that consumers can easily store). The pull-forward effect results in a fall in demand to below baseline levels after the promotion has ended (as shown below), and is something we need to understand and consider when planning future promotions:



When we're planning our future promotions for Mary's Market, one of our goals is to maximize the "gain" from competitor products (so that consumers switch from a competing toothpaste to our own) and minimize the "loss" from our own product range. If an increase in promotional sales comes at the cost of another one of our products, this is called *cannibalization,* which means that consumers have switched from one of our products they regularly buy, to the one on promotion. Some common types of cannibalization include:

- *Pack-size cannibalization*: which affects sales from other pack sizes. For example, our toothpaste might be available in 100g and 175g variants, with the larger version being more economical on a per gram basis. However, if the 100g version is placed on promotion, and the larger is not, it may result in the 100g version being more economical. We would then expect sales of the 175g version to be significantly reduced during such promotional periods because consumers will "switch" to the 100g version (thereby cannibalizing sales of the 175g version), as shown below:

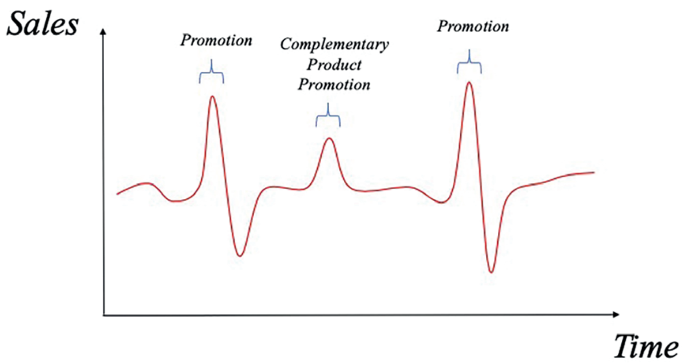


- *Sub-category cannibalization*: which affects similar products within the same product range. If we offer multiple types of whitening toothpaste, at different levels of "premium-ness" and different price points, then a promotion for premium whitening toothpaste may bring the price down on par with the mid-range version, thereby cannibalizing sales of the mid-range version.
- *Cross-category cannibalization*: which affects a different product category altogether. This is unlikely for toothpaste, but may happen for other products. For example, chocolate may cannibalize sales from biscuits, as some consumers are looking for a general snack or dessert, rather than chocolate in particular.
- *Cross-retailer cannibalization*: which occurs when consumers decide to visit a retailer where a specific product is on promotion. This type of cannibalization tends to occur with big-ticket items (e.g. television sets, coffee makers, etc.) and less so with FMCG products, however, there can still be some effect.
- *Delayed cannibalization*: is more prevalent in "treat yourself" categories, and happens when a bad-tasting product is put on special and has a delayed impact on the entire category, which could take multiple promotions to be realized. The theory being that consumers will only buy the next product from that category after finishing the one they bought, and if that product was no good, it will slow down their consumption and delay their next purchase from that category.

Through data analysis, we can create a *cannibalization matrix*, which is a table that outlines the expected cannibalization effect of certain products when placed on promotion. However, the practical challenges of constructing such a table are significant. Without applying any domain knowledge or business rules, it might be necessary to look up each individual product and calculate its cannibalizing effect on every other product in our range. For most manufacturers, which may sell hundreds or even thousands of products, this would result in a table with tens or even *hundreds of thousands of values.* While it's possible to calculate this automatically, there's no easy way to validate these values without going through them line by line.

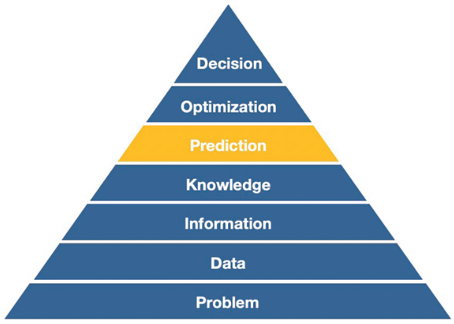
A better approach is through the application of human knowledge along with various Artificial Intelligence methods; for example, by implementing human rules based on well-founded assumptions, such as cannibalization occurring across individual categories, we could capture 80% of the expected cannibalization effect with just 20% of the modeling effort, and then complement the cannibalization matrix by additional modifications that result from a deeper analysis of the data using AI algorithms.

There's also the opposite effect of cannibalization, where an increase in sales occurs in products that aren't on promotion, but have a complementary relationship with the promoted product (in effect, creating a *cross-sell* between one product and another, even though they aren't bundled or offered together). As an example, toothpaste promotions might cause a small but measurable uplift in sales for toothbrushes, whitening kits, and mouthwash, even if those products aren't part of the promotion. This relationship is even stronger for products that are consumed together, with the classic example being pasta and pasta sauce, where pasta sales data will reveal distinct peaks when pasta sauce is on promotion, as shown below:



And finally, an increase in sales could be due to external factors; for example, sporting events can increase demand for beer and snacks, while hot weather is positively correlated with higher ice-cream sales. Not only is this type of knowledge important for decision making, but also forms the basis of our predictive modeling efforts.

Prediction



As discussed above, there are many demand drivers we need to understand within the knowledge layer, including seasonality, the correlation between variables, discount elasticity, different types of cannibalization, pull-forward effect, and so on. Furthermore, there may be long-term trends at play within the product, brand, or entire category, and we can use our knowledge of these demand drivers to develop a prediction model, which brings us to the next layer of the problem-to-decision pyramid.

Prediction models use past data to make forward-looking predictions, in effect answering the question: *Based upon what we know about the past, what's likely to happen in the future*? Prediction models can be as simple as moving average models with one input variable (historical sales), or as complex as Machine Learning models with hundreds of input variables using algorithmic methods such as random forests or neural networks (which we'll discuss in Chapter 5). Irrespective of the algorithmic method used, the goal is to achieve the highest possible accuracy. For example, after testing various algorithmic methods on historical data, we might find that neural networks provide superior accuracy when it comes to predicting the outcome of promotional plans. But after further investigation and experimentation, we might find that we can improve our accuracy even further by combining several algorithmic methods together (through an approach known as *ensemble modeling,* which we'll also explore in Chapter 5.6).

Once we've created a prediction model based on whichever algorithmic method provides us with the highest accuracy, we can then take a promotional plan (let's call it *Plan A*):

Retailer:	Mary's Market
State:	NSW
Category:	Snacks

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	…	WK 51	WK 52
Product 1	Y	y		y	Y				
Product 2	y		y		y				y
Product 3	Y		y		y	Y			y
Product 4			Y			y			
Product 5	Y			Y					
Product 6	Y	y		y					
…									
Product 100	Y			Y	Y			Y	Y

Promotion Type	In store
Promotional Price	$19.00
Shelf Price	$23.50
Discount	$4.50
Margin	35%
Customer Margin	33%
Min/Max Frequencies	13/26 Weeks
Min/Max Sell Price	$12/20
Min/Max Promo Gap	2/6 Weeks

and use our predictive model to "evaluate" this plan by predicting an outcome, such as:

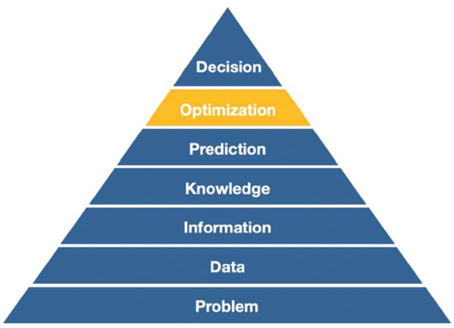
Volume	62,685
Net revenue	$1,034,303
Retailer gross profit	$227,547
Manufacturer gross profit	$268,919

In the table above, a simplified KPI report allows us to compare the effectiveness of *Plan A* against any other plan through the following metrics:

- *Volume*: which is the total unit quantity of products sold in whatever measure is used (e.g. individual units, cases, pallets, liters, etc.)
- *Net revenue*: which is the total revenue generated, based on the promotional sales price and volume
- *Retailer gross profit*: which is the total retailer gross profit, typically calculated as the difference between the promotional price and the retailer's cost of the product (plus any financial contribution the retailer made towards the promotion), multiplied by the total units sold
- *Manufacturer gross profit*: which is our gross profit, typically calculated as the difference between the net revenue and total product cost (including all production and freight costs, as well as any promotional co-funding costs)

If we implement *Plan A* for Mary's Market, then the predicted outcome is 62,685 units sold, $1,034,303 in net revenue, $227,547 in retailer profit, and $268,919 in gross profit. The accuracy of these predictions is of the utmost significance and highly dependent on the quality of our prediction model. Thus, the process of building and training a prediction model is far from trivial, and includes the careful preparation and analysis of data, as well as selecting the best algorithmic method for explaining the variability in question and producing consistent results. We'll return to these topics of data preparation and model building in Chapters 4 and 5.

Optimization



Once we've developed a prediction model for evaluating our promotional plans, we then need to create a number of plans and find the "best" one through a process of optimization. First, however, we need to define what "best" means to us, which in this case might be "any promotional plan that maximizes overall volume while satisfying our business rules and constraints." Note that some of these business rules and constraints might be for the entire category, while others are just for individual products. In the snacks category, as an example, we may have the following business rules and constraints:

- No less than 30 products and no more than 60 products on promotion in any given promotional period (soft)
- The overall minimum net revenue should be $1,000,000 (hard)
- The overall gross profit growth over last year should be 3% (hard)
- The overall minimum retailer margin growth over the last year should be 2% (hard)

whereas for Product 43, we may have some additional business rules and constraints:

- Minimum promotional price of $4.00 and maximum price of $7.00 (soft)
- Price or discount step: $0.25 (soft)
- Minimum of five and maximum of eight promotional frequencies (hard)
- Maximum of three consecutive promotional periods (hard)
- Maximum of five consecutive non-promotional periods (hard)

Such business rules and constraints typically reside in the minds of human experts within each organization, and it's often a significant undertaking to extract and document them; however, such a process is highly beneficial, because it reduces key man risk, provides visibility of the rules and constraints under which decisions are made, and allows for "testing" of each rule and constraint to ensure ongoing relevance. Recall also that these business rules and constraints can be either "hard" (meaning they cannot be violated under any circumstances) or "soft" (meaning they can be violated, but it's undesirable to do so). We can then define each constraint and/or each business rule (whether hard or soft) within a table, as follows:

	Type
Min number of products on promotion	Soft
Max number of products on promotion	Soft
Min net revenue	Hard
Min retailer margin YOY growth	Hard
Gross profit YOY growth	Hard

and do the same for each product or product category:

	Type
Min freq	Soft
Max freq	Soft
Max promo length	Soft
Min promo gap	Soft
Max promo gap	Soft
Min promo price	Hard
Max promo price	Hard
Price step	Hard
Min promos p/ period	Soft
Max promos p/ period	Soft

In some cases, violating a soft business rule may result in a better overall plan. For this reason, we must apply a "penalty" to such violations, otherwise these soft rules would always be violated and cease to be rules. Also, the penalty for some soft rules can be weighted differently to others, and we'll discuss topic of applying penalties to business rules and constraints in more detail in Chapter 6.8.

There are other important considerations related to optimization. Usually we start the search for the best plan from some starting position, for example, last year's plan or a new promotional plan that we manually create. Sometimes it's desirable to restrict the type of changes made to the original plan (e.g. whether the optimization algorithm can switch between two products, removing one from promotion and adding another) or restrict the number of changes—perhaps due to retailer requirements, or to assist in software adoption and implementation, as people can become disheartened if the slotting board they've worked on for three days comes back with 150 changes! We'll discuss these issues further in Chapter 6.9.

Now that we've defined optimality as volume maximization while satisfying our business rules and constraints, we can create many promotional plans and use our prediction model to evaluate them. At the start of this optimization process, let's say we create just two promotional plans—*Plan A* (presented earlier) and *Plan B*—and then use our prediction model to evaluate each one, with the following results:

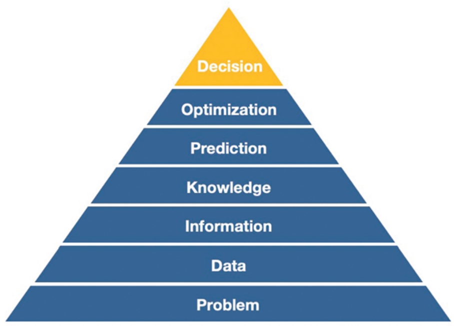
Plan A	
Volume	62,685
Net revenue	$1,034,303
Retailer gross profit	$227,547
Manufacturer gross profit	$268,919

Plan B	
Volume	59,633
Net revenue	$1,038,027
Retailer gross profit	$342,549
Manufacturer gross profit	$363,309

Assuming both plans satisfy all business rules and constraints, then *Plan A* is clearly better than *Plan B* because it has a higher predicted volume (notwithstanding that *Plan B* generates better results for the other three measures—something we'll discuss in the next section). Whether manually or through a Decision Optimization System, this process could then be repeated in a systematic way until all possible combinations are considered—the so-called *brute force approach*—if it weren't for the astronomical number of possible plans we'd have to consider, deeming such an approach impossible on even the world's fastest supercomputers (as discussed in Chapter 2.1). For this reason, "smart" algorithms are required for complex business problems such as promotional planning, and this is where Artificial Intelligence algorithms can create the most value.

And lastly, as we can see, the optimization process is relatively straightforward when we only have one objective, such as volume or revenue. To maximize this single objective, a Decision Optimization System would take an existing plan, make changes, evaluate the new plan, and continue this process in a "smart" way until the objective is maximized. This process, however, isn't so straightforward when we introduce multiple objectives.

Decision



The final decision for our promotional planning and pricing problem becomes more complicated if we need to consider multiple objectives simultaneously. This type of optimization problem is *multi-objective,* because there's more than one objective to optimize, and may result in situations where an increase in one objective results in a decrease in another.

If we return to the predicted output of *Plan A* and *Plan B* from the previous section:

Plan A	
Volume	62,685
Net revenue	$1,034,303
Retailer gross profit	$227,547
Manufacturer gross profit	$268,919

Plan B	
Volume	59,633
Net revenue	$1,038,027
Retailer gross profit	$342,549
Manufacturer gross profit	$363,309

we can see that *Plan A* has a higher predicted volume (in aggregate, across all products for the promotional period), but *Plan B* has a higher predicted total net revenue (again, in aggregate, across all products for the promotional period). If we start considering both objectives, then which promotional plan is "better"? Which one should we implement as our final decision?

In such cases, it's no longer possible to choose *the optimal* plan. Instead, we must consider a set of plans, all optimal, with some being better on one objective, while others on another. In other words, there's a trade-off between these plans, all of which are optimal in their own way. Our goal is to *understand* these trade-offs and then make the best decision. For example, a reduction in revenue should result in an increase in volume and vice versa. It's also possible to plot these promotional plans on a Pareto front (which was presented in Chapter 2.2), where all plans on the curve are optimal, meaning that it's impossible to improve any plan on one objective without suffering a decrease on some other objective. We'll also discuss multi-objective optimization problems further in Chapter 6.9.

Another complication in our decision-making process for promotional planning—even if only have a single objective—is evaluating plans where some soft business rules have been violated. As an example, let's say we create two new plans, called *Plan C*:

Retailer:	Mary's Market
State:	NSW
Category:	Snacks

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	...	WK 51	WK 52
Product 1	Y	y		y	Y				
Product 2			y						y
Product 3	Y				Y				
Product 4			Y	Y		y		Y	Y
Product 5				Y					
Product 6		y				Y		Y	Y
...									
Product 100	Y			Y	Y				Y

Promotion Type	In store
Promotional Price	$18.00
Shelf Price	$23.50
Discount	$4.25
Margin	33%
Retailer Margin	31%
Min/Max Frequencies	13/26 Weeks
Min/Max Sell Price	$12/20
Min/Max Promo Gap	2/6 Weeks

and *Plan D*:

Retailer:	Mary's Market
State:	NSW
Category:	Snacks

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	...	WK 51	WK 52
Product 1	Y			y					Y
Product 2	y		y		y				y
Product 3	Y		y		y			Y	
Product 4			Y			y		Y	
Product 5	Y			Y					
Product 6	Y	y		y	Y			Y	Y
...									
Product 100	Y	Y		Y	Y			Y	Y

Promotion Type	In store
Promotional Price	$17.50
Shelf Price	$23.50
Discount	$4.00
Margin	31%
Retailer Margin	29%
Min/Max Frequencies	13/26 Weeks
Min/Max Sell Price	$12/20
Min/Max Promo Gap	2/6 Weeks

Let's also assume that the only objective we care about is volume, and *Plan C* has a predicted volume of 61,245, while *Plan D* has a predicted volume of 60,985. So, *Plan C* is better out of these two, right? Well, not necessarily.

Both plans might violate some soft business rules, but the better plan, *Plan C*, violates more of these business rules. For example, 63 products are on promotion during Week 17 in *Plan C*, whereas one of the soft rules states that no more than 60 products should be on promotion during any given period. Similar violations occur during Week 23, 38, 39, and 48. On the other hand, there are no violations of this business rule in *Plan D*. Should we still consider *Plan C* as the better plan?

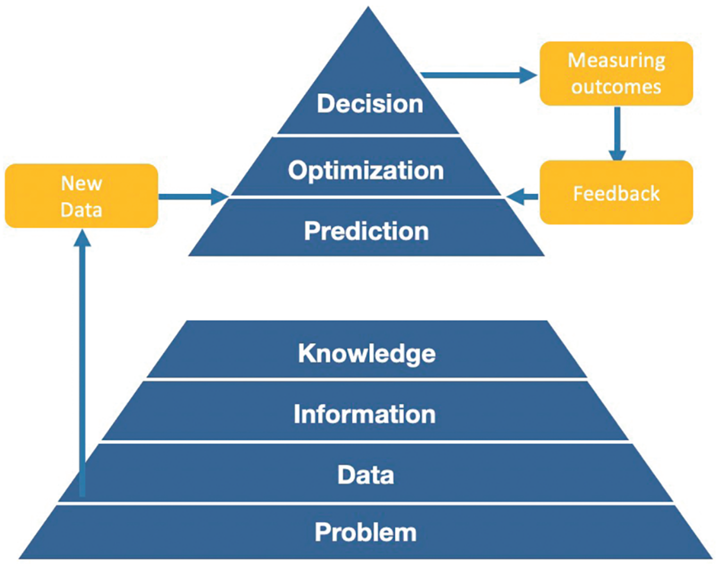
To answer this question, we need to figure out how much each violation of the min/max number of promoted products during a week is "worth," and then apply a "penalty" to the plan. For example, should each violation of the business rule be worth 100 units of volume? More? Less? Also, should we consider the degree of violation? After all, both plans violate the business rule of promoting no more than 60 products during any given period, so what's the difference between 62 products and 63? If this difference is significant, then we should set a penalty that doubles with each additional product added past 60, so exceeding this business rule by a few products would result in a small penalty in comparison to exceeding it by many products. Once we've defined these penalties for our soft business rules and constraints, we can then check the plan for violations and apply the appropriate penalty. Clearly, comparing two plans in the presence of many soft rules and constraints is far from trivial, even when optimizing on a single objective.

Adaptability

Now we can see just how complex promotional planning really is! This process has all the characteristics of a complex business problem, where the number of possible solutions is astronomical, our business rules and constraints differ from retailer to retailer, category to category, and even product to product, and our objectives and KPIs often compete with one another. Underpinning this complexity is also the fundamental challenge of predicting marketplace outcomes, which has to be done as accurately as possible.

Furthermore, all of this inherent complexity exists even if we wanted to create just one promotional plan, just once. But of course, promotional planning isn't done just once, it's an ongoing process with decisions being made in a dynamic environment where the unexpected might happen (as the COVID-19 virus has recently shown). Amidst all these planning decisions, many variables are also in a state of constant flux. New products are launched, and old ones discontinued, retailers open and shut stores, supply chain costs vary, business rules are revised, shelf space and product facings change, brands grow and decline, and consumer preferences and tastes evolve.

That's why the problem-to-decision pyramid introduces the concept of *adaptability*, which leverages a feedback loop and new data to learn from the outcome of past decisions, so that we can make even better decisions in the future:



Any prediction model that's built on historical data and never updated will gradually lose its relevance over time. To maintain accuracy and relevance, the model needs to be updated with new data, and this process usually takes place on a regular but discrete basis—perhaps monthly or weekly, depending on the environment and degree of change from one time period to the next. Furthermore, it may be necessary for the model to "forget" old data that's no longer relevant, forming a moving window of data that's used for generating predictions. We'll revisit the topic of adaptability in Chapter 7, where we'll provide a more in-depth explanation.

To see a visual example of prediction, optimization, and learning within the context of promotional planning, please watch the supplementary video for this chapter at: www.Complexica.com/RiseofAI/Chapter3.

3.3 Competitor Aspects of Promotional Planning

In this chapter, we emphasized the importance of being able to evaluate a promotional plan by accurately predicting its performance in the marketplace. However, without taking into account the strategies and corresponding actions of competitors, our prediction model might produce inaccurate results; which brings us to an entire class of real-world problems where the effectiveness of our strategy is dependent on the (yet unknown) strategies of our competitors.

Interestingly, there are parallels between trying to find the best strategy for a business problem and that of finding the best strategy for a game. Within any game, there are clear rules for what moves can be made, well-defined objectives (e.g. the criterion for winning and losing), and knowledge of who the opponents are and what moves they can make and when. Although real-world situations are far more complex, with unclear rules, many competitors (opponents), and irregular decisions (moves), many similarities exist between these two environments of a game versus the real world. In both situations, we have to devise a strategy for our moves, our opponents' countermoves, as well as decision criteria for choosing one move over another. Furthermore, the process of learning is also based on trial and error in both games and the real world.

Finding the "best" strategy (which means a strategy that is most effective against the strategy of the other player) isn't straightforward, but several algorithmic methods are available for this problem. Before describing one such method, let's first explore some *co-evolutionary* processes that exist in nature to gain a better understanding of the natural system upon which these Artificial Intelligence algorithms are based.

Given that most animals in the wild face the constant challenge of survival, their defensive and offensive survival strategies are genetically hardwired as instinctive behaviors. Some species use coloration to blend into the background, and their strategy is to remain unseen. Other species have developed a strategy based on safety in numbers, while others have learned to seek out high elevations and position themselves in a ring looking outwards, thus providing the earliest possible sighting of a potential predator. These complex strategies emerged over many generations of trial and error and illustrate the process of *co-evolution*—which is not the case of one animal against its environment, but rather an entire species against other species, each competing for resources in an environment that poses its own hostile conditions without caring about which individual animals win or lose in their struggle for existence. Competing species use random variation and selection to seek out superior survival strategies that will give them an edge over their opponents. For example, through evolution and natural selection, antelopes might improve their speed and alertness over time, with slower and less alert antelopes being eaten by lions and their genes being removed from the antelope gene pool. On the other hand, lions might become cleverer over time, so they become better hunters and catch these ever faster and more alert antelopes, with the less clever lions starving to death and their genes being removed from the lion gene pool. Hence, each "innovation" from one side may lead to an innovation from the other, which is similar to an "arms race" of inventions.

We can artificially simulate these natural co-evolutionary processes through the use of *evolutionary algorithms*, which are a type of AI optimization algorithm (covered in more detail in Chapter 6.7). The premise behind this algorithmic method is to run two optimization processes in parallel, where one represents our strategy, and the other represents the strategy of another player. The evolutionary algorithm then simulates our strategy against that of the other player and evaluates the outcome. Strategies that are more effective against the strategy of the other player are then selected, "evolved" further (i.e. optimized further), until an optimized strategy is found for the game we're playing.

In the context of promotional planning, many competing organizations face a similar problem of finding the best promotion strategy to fulfill their objectives, whether maximizing volume, revenue, market share, or some other metric. Of course, the promotional strategies of other companies aren't well known, and because of that, we can't accurately measure the impact of those strategies on our promotions. For instance, if we promote a particular product nationwide during the fifth and sixth week of the year, and one of our competitors runs a promotion during the same timeframe for a similar product, then our sales volume predictions are likely to be inaccurate.

To explore this example further, say we're concerned with the promotional strategy of our major competitor. We know the number of promotions they ran last year, and how those promotions broke down into different products and promotional periods. Using this data, we can construct a "similarity matrix" that groups our products together with those of our competitor on the basis of similarity, so that we can measure the impact on our promotions—in other words, the matrix shows which of our products were negatively impacted by our competitor's promotions of similar products last year:

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	…	WK 51	WK 52
Product Category 1	Y		Y		Y			Y	
Product Category 2		Y							
Product Category 3	Y					Y			
Product Category 4			Y					Y	
Product Category 5	Y			Y		Y			
Product Category 6	Y				Y				Y
Product Category 7					Y				
Product Category 8					Y	Y			Y

The interpretation of this "impact table" is straightforward: Some products from Product Category 1 were negatively impacted by competitor promotions of similar products during Week 1. Thus, because of this impact, we need to lower our original sales volume prediction for those products. We can go deeper, of course, improving our predictions even further by modeling promotion types and competitor pricing, but this level of detail is sufficient for now.

Even though this impact table represents our competitor's promotional strategy from last year and is unlikely to be repeated, we can use it as a starting point for the co-evolutionary process—in effect, by starting the process with our promotional plan for next year and our competitor's actual strategy from last year. By grouping all products into eight separate categories, as above, we can then model our competitor's future strategy and the impact it will have on our future promotional plan in the following way: Using evolutionary algorithms, we run two optimization processes in parallel, where the first process attempts to optimize our promotional plan (i.e. maximize volume), while the other optimization process attempts to optimize our competitor's strategy (as represented by the impact table—the better the competitor's strategy, the more impact it will have on our promotional plan). The two optimization processes then compete against each other, mimicking real co-evolutionary processes found in nature, with each process trying to "outdo" the other. In other words, one optimization process is striving to create a promotional plan that maximizes sales volume while taking into account the most damaging impact table from the other optimization process; while the other optimization process is striving to create an impact table that maximizes damage to our best promotional plan.

The "connector" between these two optimization processes is the evaluation of each plan and impact table. For example, to evaluate our promotional plan, we need to know the best impact table from the other optimization process, and then lower our volume prediction accordingly—and vice versa, to evaluate an impact table, our competitor needs to know how much damage their impact table caused to our best promotional plan. Let's illustrate this back and forth process by re-visiting *Plan A* from above:

Retailer:	Mary's Market
State:	NSW
Category:	Snacks

	WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	...	WK 51	WK 52
Product 1	Y	y		y	Y				
Product 2	y		y		y				y
Product 3	Y		y		y	Y			y
Product 4			Y			y			
Product 5	Y			Y					
Product 6	Y	y		y					
...									
Product 100	Y			Y	Y			Y	Y

Promotion Type	In store
Promotional Price	$19.00
Shelf Price	$23.50
Discount	$4.50
Margin	35%
Customer Margin	33%
Min/Max Frequencies	13/26 Weeks
Min/Max Sell Price	$12/20
Min/Max Promo Gap	2/6 Weeks

To evaluate *Plan A*, we can take the best impact table from the second optimization process—as this table represents the best (current) strategy of our competitor—and then use it to measure the impact on our promotional plan and reduce our sales volume predictions accordingly. Because we know our best promotional plan at each moment of the optimization process, we can estimate the quality of our competitor's strategy on the basis of the damage it does to our promotional plan—the larger the damage, the better their strategy. In other words, we can take into account the best strategy of our competitor whenever evaluating our own promotional plan. This means that our promotional plans become better over time, against the superior strategies of our competitor! Thus, the co-evolutionary process may provide us with not only the best promotional plan, but also insights into the likely behavior of our competition.

And finally, co-evolutionary algorithms aren't limited to just one competitor. If we have three primary competitors, we can introduce three additional "players" into the "game," where each player runs a process of optimizing its strategy against those of the other players. To evaluate a promotional plan in this situation, we would take the best strategies from the other three optimization processes, which represent the best current strategies of our competitors. This allows us to improve the effectiveness and robustness of our promotional plan by taking into account the best possible strategy of our competitors, as well as upgrading our prediction model (which was originally based only on our promotional plan) by incorporating data from our competitor's impact table.

The use of co-evolutionary methods is a great example not only of Artificial Intelligence algorithms, but also of how AI scientists go about replicating natural processes in order to solve problems of great complexity (in much the same way that the natural behavior of ants is artificially replicated in order to create ant algorithms, as discussed in Chapter 1.3). In the next part of this book we'll discuss how AI algorithms are used for prediction, optimization, and learning, but because data represents a key part of the whole process, we'll begin with a chapter on data and modeling. For more information on the material covered in this chapter, please watch the supplementary video at: www.Complexica.com/RiseofAI/Chapter3.